

I. GÉNÉRALITÉS SUR LES MATRICES SYMÉTRIQUES RÉELLES

1 D'après le théorème spectral, toute matrice symétrique réelle est orthodiagonalisable. Réciproquement, si $A \in \mathcal{M}_n(\mathbb{R})$ est orthodiagonalisable, il existe une matrice diagonale D et une matrice orthogonale P telles que $A = PDP^\top$. Comme la matrice D est diagonale, elle est symétrique, donc $D^\top = D$. Il vient

$$A^\top = (PDP^\top)^\top = (P^\top)^\top D^\top P^\top = PDP^\top = A$$

Finalement,

Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est orthodiagonalisable si, et seulement si, elle est symétrique.

2 Remarquons que la somme des première et troisième colonnes de A_1 vaut

$$\begin{pmatrix} 7 & 0 & 7 \end{pmatrix}^\top = 7 \begin{pmatrix} 1 & 0 & 1 \end{pmatrix}^\top$$

On en déduit que

Le vecteur $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ est un vecteur propre de A_1 associé à la valeur propre $\lambda_1 = 7$.

3 Calculons $A_1 - \lambda_1 I_3 = A_1 - 7I_3 = \begin{pmatrix} -4 & -2 & 4 \\ -2 & -1 & 2 \\ 4 & 2 & -4 \end{pmatrix}$

La matrice $A_1 - \lambda_1 I_3$ est de rang 1 car tous ses vecteurs colonnes sont colinéaires au vecteur $\begin{pmatrix} 2 & 1 & -2 \end{pmatrix}^\top$. D'après le théorème du rang, le sous-espace propre de A_1 associé à la valeur propre λ_1 est donc de dimension 2 et c'est le plan d'équation cartésienne $2x + y - 2z = 0$. D'après la question 2, on peut prendre comme premier vecteur propre $X_1 = \begin{pmatrix} 1 & 0 & 1 \end{pmatrix}^\top$ (il correspond à $x = 1$ et $y = 0$).

Pour le deuxième vecteur propre, on peut prendre $\begin{pmatrix} 1 & -2 & 0 \end{pmatrix}^\top$ (à l'aide de $x = 1$ et $z = 0$). Mais comme la question 4 demande d'orthodiagonaliser A_1 , autant prendre directement un vecteur X_2 orthogonal à X_1 , c'est-à-dire tel que $(X_1 | X_2) = 0$. Pour cela, il suffit de résoudre le système

$$\begin{cases} 2x + y - 2z = 0 \\ x + z = 0 \end{cases} \quad \text{soit} \quad \begin{cases} 4x + y = 0 \\ x + z = 0 \end{cases} \quad \text{d'où} \quad \begin{cases} y = -4x \\ z = -x \end{cases}$$

qui mène par exemple à $X_2 = \begin{pmatrix} -1 & 4 & 1 \end{pmatrix}^\top$ (avec $x = -1$, qui implique $y = 4$ et $z = 1$). La famille (X_1, X_2) est bien une base de l'espace propre E_{λ_1} car c'est une famille orthogonale de deux vecteurs non nuls dans un espace de dimension 2. Ainsi,

$$E_{\lambda_1}(A_1) = \text{Vect}(X_1, X_2) \quad \text{avec} \quad X_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{et} \quad X_2 = \begin{pmatrix} -1 \\ 4 \\ 1 \end{pmatrix}$$

Comme la matrice A_1 est diagonalisable, la somme de ses valeurs propres est égale à sa trace, donc, en notant λ_2 la seconde valeur propre de A_1 ,

$$\lambda_1 + \lambda_1 + \lambda_2 = \text{Tr}(A_1) = 3 + 6 + 3 = 12$$

Comme $\lambda_1 = 7$, on en déduit que $\lambda_2 = -2$. Par suite,

$$\text{sp}(A_1) = \{-2, 7\}$$

4 La matrice A_1 est symétrique réelle. Elle est donc orthodiagonalisable d'après la question 1. La question 3 affirme que $\lambda_1 = 7$, $\lambda_2 = -2$ et que les sous-espaces propres associés sont de dimension respective 2 et 1. Comme les sous-espaces propres de A_1 sont orthogonaux, on peut chercher un vecteur propre $X_3 = (x \ y \ z)^\top$ associé à la valeur propre λ_2 en résolvant le système

$$\begin{cases} 0 = (X_1 | X_3) = x + z \\ 0 = (X_2 | X_3) = -x + 4y + z \end{cases} \quad \text{qui s'écrit} \quad \begin{cases} 0 = x + z \\ 0 = 4y + 2z \end{cases}$$

En prenant $y = 1$, on obtient le vecteur $X_3 = (2 \ 1 \ -2)^\top$.

On pouvait aussi déterminer un vecteur propre X_3 associé à la valeur propre λ_2 de deux autres manières. D'abord, en exploitant l'orthogonalité des sous-espaces propres comme ci-dessus, on pouvait calculer directement le produit vectoriel entre X_1 et X_2 :

$$X_3 = X_1 \wedge X_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \wedge \begin{pmatrix} -1 \\ 4 \\ 1 \end{pmatrix} = \begin{pmatrix} -4 \\ -2 \\ 4 \end{pmatrix}$$

Sinon, on pouvait aussi calculer

$$A_1 - \lambda_2 I_3 = A_1 + 2I_3 = \begin{pmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{pmatrix}$$

et chercher les vecteurs propres de A_1 associés à λ_2 en résolvant le système

$$\begin{cases} 5x - 2y + 4z = 0 \\ -2x + 8y + 2z = 0 \\ 4x + 2y + 5z = 0 \end{cases}$$

On dispose donc d'une base orthogonale de vecteurs propres (X_1, X_2, X_3) . Il reste à normer ces trois vecteurs pour obtenir une base orthonormée de vecteurs propres, en calculant

$$\|X_1\| = \sqrt{1^2 + 0^2 + 1^2} = \sqrt{2} \quad \|X_2\| = \sqrt{(-1)^2 + 4^2 + 1^2} = \sqrt{18} = 3\sqrt{2}$$

et

$$\|X_3\| = \sqrt{2^2 + 1^2 + (-2)^2} = 3$$

Il vient

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{2}{3} \\ 0 & \frac{4}{3\sqrt{2}} & \frac{1}{3} \\ \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & -\frac{2}{3} \end{pmatrix} = \frac{1}{3\sqrt{2}} \begin{pmatrix} 3 & -1 & 2\sqrt{2} \\ 0 & 4 & \sqrt{2} \\ 3 & 1 & -2\sqrt{2} \end{pmatrix}$$

$$A_1 = PDP^\top \quad \text{avec} \quad D = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{pmatrix} \quad \text{et} \quad P = \frac{1}{3\sqrt{2}} \begin{pmatrix} 3 & -1 & 2\sqrt{2} \\ 0 & 4 & \sqrt{2} \\ 3 & 1 & -2\sqrt{2} \end{pmatrix}$$

5 Montrons que l'application ϕ est bilinéaire symétrique définie positive. Considérons $(P, Q, R) \in (\mathbb{R}_{n-1}[X])^3$ et $\lambda \in \mathbb{R}$.

- L'application ϕ est symétrique puisque

$$\phi(P, Q) = \int_0^1 P(t)Q(t) dt = \int_0^1 Q(t)P(t) dt = \phi(Q, P)$$

- D'après la linéarité de l'intégrale, on a

$$\begin{aligned}\phi(\lambda P + R, Q) &= \int_0^1 [\lambda P(t) + R(t)] Q(t) dt \\ &= \lambda \int_0^1 P(t) Q(t) dt + \int_0^1 R(t) Q(t) dt \\ \phi(\lambda P + R, Q) &= \lambda \phi(P, Q) + \phi(R, Q)\end{aligned}$$

ce qui montre la linéarité à gauche de ϕ , donc sa bilinéarité par symétrie.

- Comme P est à coefficients réels, $P(t)^2 \geq 0$ pour tout $t \in [0; 1]$ et la positivité de l'intégrale donne

$$\phi(P, P) = \int_0^1 P(t)^2 dt \geq 0$$

donc l'application ϕ est positive.

- Enfin, si $\phi(P, P) = 0$, alors la fonction polynomiale $t \mapsto P(t)^2$ est continue et d'intégrale nulle sur $[0; 1]$. Elle est donc nulle d'après la positivité stricte de l'intégrale. Par conséquent, le polynôme P admet une infinité de racines, si bien que $P = 0$. Ceci prouve que ϕ est définie.

L'application ϕ est un produit scalaire sur $\mathbb{R}_{n-1}[X]$.

- 6** | Attention, à partir de cette question, les indices des lignes et des colonnes des matrices vont de 0 à $n - 1$ (ce qui est usuel en Python par exemple), et non de 1 à n comme on en a l'habitude.

Soit $(i, j) \in \llbracket 0; n - 1 \rrbracket^2$. On a

$$h_{i,j} = \phi(X^i, X^j) = \int_0^1 t^i t^j dt = \left[\frac{1}{i+j+1} t^{i+j+1} \right]_0^1 = \frac{1}{i+j+1}$$

La matrice H du produit scalaire ϕ dans la base canonique de $\mathbb{R}_{n-1}[X]$ est égale à

$$H = (h_{i,j})_{i,j \in \{0, \dots, n-1\}} = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \ddots & \ddots & \frac{1}{n+1} \\ \frac{1}{3} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \frac{1}{2n-2} \\ \frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n-2} & \frac{1}{2n-1} \end{pmatrix}$$

- 7** | Notons $U = (u_0 \quad \dots \quad u_{n-1})^\top \in \mathcal{M}_{n,1}(\mathbb{R})$. Soit $i \in \llbracket 0; n - 1 \rrbracket$. On a

$$(HU)_i = \sum_{j=0}^{n-1} h_{i,j} u_j \quad \text{puis} \quad U^\top H U = \sum_{i=0}^{n-1} u_i (HU)_i = \sum_{i=0}^{n-1} u_i \sum_{j=0}^{n-1} h_{i,j} u_j$$

Finalement,

$$\forall U \in \mathcal{M}_{n,1}(\mathbb{R}) \quad U^\top H U = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \phi(X^i, X^j) u_i u_j$$

Remarquons tout d'abord que la valeur encadrée ci-dessus vaut aussi

$$\phi\left(\sum_{i=0}^{n-1} u_i X_i, \sum_{j=0}^{n-1} u_j X_j\right) = \left\| \sum_{i=0}^{n-1} u_i X_i \right\|^2$$

Cette expression sera utile dans la question 8.

En outre, comme ce calcul servira à plusieurs reprises dans ce problème, donnons la valeur du réel $X^\top AY$ dans le cas général où $A \in \mathcal{M}_n(\mathbb{R})$ et $(X, Y) \in (\mathcal{M}_{n,1}(\mathbb{R}))^2$ en notant comme ci-dessus $A = (a_{i,j})_{0 \leq i,j \leq n-1}$, $X = (x_i)_{0 \leq i \leq n-1}$ et $Y = (y_i)_{0 \leq i \leq n-1}$. Soit $(i, j) \in \llbracket 0; n-1 \rrbracket^2$. Alors

$$(AY)_i = \sum_{j=0}^{n-1} a_{i,j} y_j$$

$$\text{puis } X^\top AY = \sum_{i=0}^{n-1} x_i (AY)_i = \sum_{i=0}^{n-1} x_i \sum_{j=0}^{n-1} a_{i,j} y_j = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} x_i a_{i,j} y_j$$

8 D'après la question 6, pour tout $(i, j) \in \llbracket 0; n-1 \rrbracket^2$,

$$h_{i,j} = \frac{1}{i+j+1} = \frac{1}{j+i+1} = h_{j,i}$$

donc

$$H \in \mathcal{S}_n(\mathbb{R})$$

On dit que H est la matrice du produit scalaire ϕ dans la base canonique de $\mathbb{R}_{n-1}[X]$.

Soient $\lambda \in \mathbb{R}$ une valeur propre de H et $U \in \mathcal{M}_{n,1}(\mathbb{R})$ un vecteur propre associé, de sorte que $HU = \lambda U$. D'une part,

$$U^\top HU = \lambda U^\top U = \lambda \|U\|^2$$

avec $\|U\| \neq 0$ puisque U est un vecteur propre. D'autre part, le résultat de la question précédente et la bilinéarité du produit scalaire ϕ assurent que

$$\begin{aligned} U^\top HU &= \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \phi(X^i, X^j) u_i u_j \\ &= \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \phi(u_i X^i, u_j X^j) \\ &= \phi\left(\sum_{i=0}^{n-1} u_i X^i, \sum_{j=0}^{n-1} u_j X^j\right) \\ U^\top HU &= \left\| \sum_{i=0}^{n-1} u_i X^i \right\|^2 \\ &= \frac{\left\| \sum_{i=0}^{n-1} u_i X^i \right\|^2}{\|U\|^2} > 0 \end{aligned}$$

Par suite,

$$\lambda = \frac{\left\| \sum_{i=0}^{n-1} u_i X^i \right\|^2}{\|U\|^2} > 0$$

Ceci étant vrai pour toute valeur propre λ de H ,

Les valeurs propres de H sont strictement positives.

9 Pour les questions 9 à 12, on considère une matrice $A \in \mathcal{M}_n(\mathbb{R})$ de spectre non vide.

Si A est nilpotente, il existe $p \in \mathbb{N}^*$ tel que $A^p = 0$. Ainsi, $0 = \det(A^p) = \det(A)^p$, donc $\det(A) = 0$ puis A n'est pas inversible. Il existe donc un vecteur $X \in \mathcal{M}_{n,1}(\mathbb{R})$ tel que $AX = 0$, d'où $0 \in \text{sp}(A)$.

Si $\lambda \in \mathbb{R}$ est une valeur propre de A et si $X \in \mathcal{M}_{n,1}(\mathbb{R})$ est un vecteur propre associé, alors $AX = \lambda X$, puis $A^k X = \lambda^k X$ pour tout $k \geq 1$ par récurrence immédiate. En particulier, $0 = A^p X = \lambda^p X$. Comme X est non nul, on en déduit que $\lambda^p = 0$ puis que $\lambda = 0$. Par conséquent, $\text{sp}(A) \subset \{0\}$.

Par suite, $\text{sp}(A) = \{0\}$. Finalement,

Si A est nilpotente, alors son rayon spectral $\rho(A)$ est nul.

10 Remarquons que $C = \{U \in \mathcal{M}_{n,1}(\mathbb{R}) \mid \|U\| = 1\}$ est la sphère de centre 0 et de rayon 1 associée à la norme euclidienne, donc d'après le cours,

L'espace C est une partie fermée de $\mathcal{M}_{n,1}(\mathbb{R})$.

11 L'application $U \mapsto |U^\top AU|$ est continue en tant que composition des applications $x \mapsto |x|$ continue sur $\mathbb{R}$. De même, l'application $(U, V) \mapsto U^\top V$ est continue sur $\mathcal{M}_{n,1}(\mathbb{R})^2$ car elle est bilinéaire en dimension finie. Enfin, l'application $U \mapsto AU$ est continue sur $\mathcal{M}_{n,1}(\mathbb{R})$ car elle est linéaire en dimension finie. De plus, d'après la question précédente, l'espace C est une partie fermée bornée (par 1) de $\mathcal{M}_{n,1}(\mathbb{R})$. L'application $U \mapsto |U^\top AU|$ est donc bornée sur C et elle atteint ses bornes sur C (d'après le théorème des bornes atteintes), ce qui signifie en particulier qu'elle possède un maximum.

L'application $U \mapsto |U^\top AU|$ admet un maximum sur C .

12 Soit λ une valeur propre de A telle que $\rho(A) = |\lambda|$ et X un vecteur propre associé. Quitte à le normer puisqu'il est non nul, on peut le supposer unitaire, donc appartenant à la partie C . Ainsi, $U \in C$ et

$$|U^\top AU| = |U^\top \lambda U| = |\lambda| \|U\|^2 = |\lambda| = \rho(A)$$

On en déduit que

$$\rho(A) \leq \max_{U \in C} |U^\top AU|$$

On sait que le spectre réel de $A \in \mathcal{M}_n(\mathbb{R})$ est non vide et que c'est un ensemble fini de réels. Ceci justifie que $\rho(A)$ est un maximum (nécessairement atteint en une valeur) et pas seulement une borne supérieure.

13 Puisque $A \in \mathcal{S}_n(\mathbb{R})$, la matrice A est orthodagonalisable d'après la question 1. Il existe donc une matrice diagonale réelle $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ et une matrice orthogonale P telles que $A = PDP^\top$. Comme le spectre de $A \in \mathcal{M}_n(\mathbb{R})$ est non vide, on a $\rho(A) \leq \max_{U \in C} |U^\top AU|$ d'après la question précédente.

Montrons à présent l'inégalité inverse. Soit $U \in C$. On a

$$U^\top AU = U^\top PDP^\top U = (P^\top U)^\top D(P^\top U) = V^\top DV$$

en posant $V = P^\top U$. Notons que

$$V^\top V = (P^\top U)^\top (P^\top U) = U^\top PP^\top U = U^\top U$$

car P est orthogonale. Comme $U \in C$, $U^\top U = 1$ donc $V^\top V = 1$, puis $V \in C$.

La matrice $P^\top$ est orthogonale, donc c'est la matrice d'un endomorphisme orthogonal, qui conserve la norme. De ce fait, $P^\top U \in C$ pour tout $U \in C$.

En écrivant $V = (v_1 \ \dots \ v_n)^\top$, un calcul similaire à celui effectué à la question 7 assure que

$$U^\top A U = V^\top D V = \sum_{i=0}^{n-1} \lambda_i v_i^2$$

$$\text{d'où} \quad |U^\top A U| \leq \sum_{i=0}^{n-1} |\lambda_i| v_i^2 \leq \rho(A) \sum_{i=0}^{n-1} v_i^2 = \rho(A) \underbrace{\|V\|^2}_{=1} = \rho(A)$$

Ceci étant vrai pour tout $U \in C$,

$$\max_{U \in C} |U^\top A U| \leq \rho(A)$$

Finalement,

$$\boxed{\rho(A) = \max_{U \in C} |U^\top A U|}$$

14 Dans le cas où les valeurs propres de A sont positives, le calcul de la question précédente montre que

$$\forall U \in C \quad U^\top A U = \sum_{i=0}^{n-1} \lambda_i v_i^2 \geq 0$$

ce qui s'écrit

$$\forall U \in C \quad U^\top A U = |U^\top A U|$$

Ainsi,

$$\boxed{\rho(A) = \max_{U \in C} (U^\top A U)}$$

15 Montrons que l'application ρ définit une norme sur $\mathcal{S}_n(\mathbb{R})$. Tout d'abord, elle est bien définie sur cet espace car les matrices symétriques réelles sont diagonalisables sur $\mathbb{R}$, elles ont donc bien un spectre réel non vide.

- Par définition, ρ est bien à valeurs dans $\mathbb{R}_+$.
- Comme la matrice nulle n'admet que 0 comme valeur propre, $\rho(0_n) = 0$. Soit $A \in \mathcal{S}_n(\mathbb{R})$ telle que $\rho(A) = 0$. Alors $\text{sp}(A) = \{0\}$. Comme la matrice A est diagonalisable, cela signifie qu'elle est nulle. Autrement dit,

$$\rho(A) = 0 \iff A = 0_n$$

- Soient $A \in \mathcal{S}_n(\mathbb{R})$ et $\mu \in \mathbb{R}$. Si μ est non nul, pour tout réel $\lambda \in \mathbb{R}$ et tout vecteur non nul $X \in \mathbb{R}^n$, on a équivalence entre les propositions

$$AX = \lambda X \iff (\mu A)X = (\mu\lambda)X$$

donc $\text{sp}(\mu A) = \mu \text{sp}(A)$. Cette égalité restant vraie pour $\mu = 0$ puisque le spectre de la matrice nulle est réduit à $\{0\}$,

$$\rho(\mu A) = \max_{\lambda \in \text{sp}(A)} |\mu\lambda| = |\mu| \max_{\lambda \in \text{sp}(A)} |\lambda| = |\mu| \rho(A)$$

- Enfin, soit $(A, B) \in (\mathcal{S}_n(\mathbb{R}))^2$. Utilisons l'expression obtenue à la question 13. Soit $U \in C$. L'inégalité triangulaire donne alors

$$|U^\top (A + B) U| \leq |U^\top A U| + |U^\top B U| \leq \rho(A) + \rho(B)$$

Par définition du maximum sur $U \in C$, il vient

$$\rho(A + B) \leq \rho(A) + \rho(B)$$

En conclusion,

$$\boxed{\text{L'application } \rho \text{ définit une norme sur } \mathcal{S}_n(\mathbb{R}).}$$

II. MATRICE DE COVARIANCE

16 À partir de cette question, les indices des lignes et des colonnes des matrices vont de 1 à n de manière plus usuelle.

Par définition, pour tout $(i, j) \in \llbracket 1; n \rrbracket^2$,

$$\sigma_{i,j} = \text{cov}(Y_i, Y_j) = \text{cov}(Y_j, Y_i) = \sigma_{j,i}$$

donc

La matrice Σ_Y est symétrique.

De plus, pour tout $(i, j) \in \llbracket 1; n \rrbracket^2$,

$$\begin{aligned}\sigma_{i,j} &= \text{cov}(Y_i, Y_j) \\ &= \mathbb{E}((Y_i - \mathbb{E}(Y_i))(Y_j - \mathbb{E}(Y_j))) \\ \sigma_{i,j} &= \left(\mathbb{E} \left((Y - \mathbb{E}(Y))(Y - \mathbb{E}(Y))^T \right) \right)_{i,j}\end{aligned}$$

Rappelons que si $(U, V) \in (\mathcal{M}_{n,1}(\mathbb{R}))^2$ et $(i, j) \in \llbracket 1; n \rrbracket^2$,

$$(UV^T)_{i,j} = U_i V_j$$

Finalement,

$$\Sigma_Y = \mathbb{E} \left((Y - \mathbb{E}(Y))(Y - \mathbb{E}(Y))^T \right)$$

Soient $U = (u_1 \ \dots \ u_n)^T \in \mathcal{M}_{n,1}(\mathbb{R})$ un vecteur constant et $(i, j) \in \llbracket 1; n \rrbracket^2$. Alors le coefficient d'indice (i, j) de la matrice de covariance Σ_{Y+U} est égal à

$$\text{cov}(Y_i + u_i, Y_j + u_j) = \text{cov}(Y_i, Y_j) + \text{cov}(Y_i, u_j) + \text{cov}(u_i, Y_j) + \text{cov}(u_i, u_j)$$

par bilinéarité de la covariance. En outre, u_i et u_j étant des variables aléatoires constantes, les trois dernières covariances sont nulles. En effet, si X et Y sont deux variables aléatoires et que X est constante, alors $X = \mathbb{E}(X)$, donc

$$\text{cov}(X, Y) = \mathbb{E} \left[\underbrace{(X - \mathbb{E}(X))}_{=0} (Y - \mathbb{E}(Y)) \right] = 0$$

et $\text{cov}(Y, X) = \text{cov}(X, Y) = 0$ par symétrie de la covariance. Finalement,

$$\forall (i, j) \in \llbracket 1; n \rrbracket^2 \quad \text{cov}(Y_i + u_i, Y_j + u_j) = \text{cov}(Y_i, Y_j)$$

soit

Si $U \in \mathcal{M}_{n,1}(\mathbb{R})$ est un vecteur constant, alors $\Sigma_{Y+U} = \Sigma_Y$.

17 Soient $p \in \mathbb{N}^*$ et $M \in \mathcal{M}_{p,n}(\mathbb{R})$. Par définition du vecteur aléatoire $Z = MY$,

$$\forall i \in \llbracket 1; p \rrbracket \quad Z_i = \sum_{j=1}^n m_{i,j} Y_j$$

Pour tout $i \in \llbracket 1; p \rrbracket$, la variable aléatoire Z_i est donc une combinaison linéaire des variables aléatoires $Y_1, \dots, Y_n$. Ces dernières admettent une espérance finie d'après l'énoncé, donc la variable Z_i admet également une espérance finie, d'où l'existence de $\mathbb{E}(Z)$. En outre, la linéarité de l'espérance assure que

$$\forall i \in \llbracket 1; p \rrbracket \quad \mathbb{E}(Z_i) = \sum_{j=1}^n m_{i,j} \mathbb{E}(Y_j)$$

ce qui s'écrit

$$\mathbb{E}(Z) = M\mathbb{E}(Y)$$

Si $p \in \mathbb{N}^*$ et $M \in \mathcal{M}_{p,n}(\mathbb{R})$, la variable aléatoire $Z = MY$ admet une espérance et on a l'égalité $\mathbb{E}(Z) = M\mathbb{E}(Y)$.

De la même manière, $\text{cov}(Y_i, Y_j)$ est bien définie pour tout $(i, j) \in \llbracket 1; n \rrbracket^2$ puisque Σ_Y existe. De ce fait, $\text{cov}(Z_i, Z_j)$ existe pour tout $(i, j) \in \llbracket 1; p \rrbracket^2$ et vaut par bilinéarité de la covariance

$$\begin{aligned}\text{cov}(Z_i, Z_j) &= \text{cov}\left(\sum_{k=1}^n m_{i,k} Y_k, \sum_{\ell=1}^n m_{j,\ell} Y_\ell\right) \\ &= \sum_{k=1}^n m_{i,k} \sum_{\ell=1}^n m_{j,\ell} \text{cov}(Y_k, Y_\ell) \\ \text{cov}(Z_i, Z_j) &= \sum_{k=1}^n \sum_{\ell=1}^n m_{i,k} \sigma_{k,\ell} m_{j,\ell} = (\mathbf{M} \Sigma_Y \mathbf{M}^\top)_{i,j}\end{aligned}$$

si bien que

$$\Sigma_Z = \mathbf{M} \Sigma_Y \mathbf{M}^\top$$

Si $p \in \mathbb{N}^*$ et $\mathbf{M} \in \mathcal{M}_{p,n}(\mathbb{R})$, la variable aléatoire $Z = \mathbf{M}Y$ admet une matrice de covariance Σ_Z et $\Sigma_Z = \mathbf{M} \Sigma_Y \mathbf{M}^\top$.

18 La matrice Σ_Y est symétrique réelle d'après la question 16, elle est donc orthodagonalisable d'après le résultat de la question 1. Avec les notations de l'énoncé, la matrice $\mathbf{P}$ est la matrice de passage de la base canonique de $\mathcal{M}_{n,1}(\mathbb{R})$ à une base orthonormée formée de vecteurs propres de Σ_Y . D'après les formules de changement de base, il existe donc une matrice $\mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_n)$ telle que

$$\Sigma_Y = \mathbf{P} \mathbf{D} \mathbf{P}^\top$$

Notons que les réels $\lambda_1, \dots, \lambda_n$ sont alors les valeurs propres de Σ_Y . Comme $\mathbf{X} = \mathbf{P}^\top \mathbf{Y}$, le résultat de la question 17 assure que la variable $\mathbf{X}$ admet une matrice de covariance Σ_X et que

$$\Sigma_X = \mathbf{P}^\top \Sigma_Y (\mathbf{P}^\top)^\top = \mathbf{P}^\top \Sigma_Y \mathbf{P}$$

En réinjectant la valeur de Σ_Y dans cette égalité, il vient

$$\Sigma_X = \mathbf{P}^\top \mathbf{P} \mathbf{D} \mathbf{P}^\top \mathbf{P} = \mathbf{D}$$

puisque $\mathbf{P}$ est une matrice orthogonale. Ainsi,

La matrice Σ_X est diagonale.

19 D'après le calcul effectué dans la question précédente, les matrices Σ_Y et Σ_X sont semblables. Elles ont donc les mêmes valeurs propres notées $(\lambda_1, \dots, \lambda_n)$ qui sont les coefficients diagonaux de Σ_X :

$$\forall i \in \llbracket 1; n \rrbracket \quad \lambda_i = (\Sigma_X)_{i,i} = \text{cov}(X_i, X_i) = \mathbb{V}(X_i) \geq 0$$

Les valeurs propres de Σ_Y sont toutes positives.

20 D'après ce qui précède, les matrices Σ_Y et Σ_X sont semblables donc leurs traces sont égales. Par définition, la variance totale de $\mathbf{X}$ est égale à

$$\mathbb{V}_T(\mathbf{X}) = \sum_{i=1}^n \mathbb{V}(X_i) = \text{Tr}(\Sigma_X) = \text{Tr}(\Sigma_Y) = \sum_{i=1}^n \mathbb{V}(Y_i) = \mathbb{V}_T(\mathbf{Y})$$

Les variances totales de $\mathbf{X}$ et de $\mathbf{Y}$ sont égales.

On pouvait aussi remarquer que les matrices Σ_Y et Σ_X sont diagonalisables et qu'elles ont les mêmes valeurs propres $(\lambda_1, \dots, \lambda_n)$. Leur trace est donc égale à la somme de ces valeurs propres :

$$\text{Tr}(\Sigma_X) = \sum_{i=1}^n \lambda_i = \text{Tr}(\Sigma_Y)$$

et on conclut comme ci-dessus.

21 Il suffit de considérer des variables aléatoires discrètes mutuellement indépendantes vérifiant, pour tout $i \in \llbracket 1; n \rrbracket$, l'égalité $\mathbb{V}(Z_i) = \lambda_i$. D'après le cours, on sait que de telles variables aléatoires existent. Par construction, on en déduit l'existence d'un vecteur aléatoire $Z = (Z_1 \dots Z_n)^\top$ tel que Σ_Z existe et $\Sigma_Z = D$.

Plusieurs exemples de lois usuelles conviennent, le plus simple étant sans doute la loi de Poisson de paramètre λ_i pour tout $i \in \llbracket 1; n \rrbracket$ puisque que l'on sait d'après le cours que la variance d'une telle loi est égale à λ_i .

Il existe une variable aléatoire discrète Z à valeurs dans $\mathcal{M}_{n,1}(\mathbb{R})$ telle que $\Sigma_Z = D$.

22 La matrice A est symétrique réelle, elle est donc orthodiagonalisable d'après le résultat de la question 1. Il existe donc une matrice diagonale D et une matrice orthogonale P telles que

$$A = PDP^\top$$

D'après le résultat de la question 21, il existe une variable aléatoire discrète Z à valeurs dans $\mathcal{M}_{n,1}(\mathbb{R})$ telle que $\Sigma_Z = D$. Posons $Y = PZ$. Alors, d'après le résultat de la question 17, la matrice de covariance Σ_Y de Y existe et

$$\Sigma_Y = P\Sigma_Z P^\top = PDP^\top = A$$

Il existe une variable aléatoire discrète Y à valeurs dans $\mathcal{M}_{n,1}(\mathbb{R})$ telle que $\Sigma_Y = A$.

23 Comme $X = U^\top Y$, utilisons la question 17 en posant $M = U^\top$. Alors X admet une matrice de covariance Σ_X vérifiant $\Sigma_X = U^\top \Sigma_Y U$. En outre, comme X est une variable aléatoire réelle, ceci prouve l'existence de la variance de X puisque $\mathbb{V}(X) = \Sigma_X$. Finalement,

La variable aléatoire X admet une variance, qui vaut $\mathbb{V}(X) = U^\top \Sigma_Y U$.

En toute rigueur, $\Sigma_X = U^\top \Sigma_Y U$ est une matrice carrée de taille 1 et $\mathbb{V}(X)$ est un réel. On se permet d'identifier ces valeurs comme dans l'énoncé.

24 Si $r = n$, alors $\text{Im } \Sigma_Y = \mathcal{M}_{n,1}(\mathbb{R})$. Comme la variable aléatoire Y est à valeurs dans $\mathcal{M}_{n,1}(\mathbb{R})$, l'événement $\{Y - \mathbb{E}(Y) \in \text{Im } \Sigma_Y\}$ est l'événement certain. Ainsi,

Si $r = n$, alors $\mathbb{P}(Y - \mathbb{E}(Y) \in \text{Im } \Sigma_Y) = 1$.

25 Soient $U \in \text{Ker } \Sigma_Y$ et $V \in \text{Im } \Sigma_Y$. Il existe $W \in \mathcal{M}_{n,1}(\mathbb{R})$ tel que $V = \Sigma_Y W$. Par définition du produit scalaire,

$$(U | V) = U^\top V = U^\top \Sigma_Y W = (\Sigma_Y^\top U)^\top W$$

Comme Σ_Y est symétrique, ceci entraîne que

$$(U | V) = (\Sigma_Y U)^\top W = 0$$

puisque $\Sigma_Y U = 0$. Par suite, les espaces $\text{Ker } \Sigma_Y$ et $\text{Im } \Sigma_Y$ sont orthogonaux. On en déduit directement que $\text{Ker } \Sigma_Y \cap \text{Im } \Sigma_Y = \{0\}$ (le vecteur nul étant le seul vecteur orthogonal à lui-même). Ces espaces sont donc en somme directe. D'après le théorème du rang, $\dim \text{Ker } \Sigma_Y + \dim \text{Im } \Sigma_Y = \dim \mathcal{M}_{n,1}(\mathbb{R})$, donc ils sont supplémentaires dans $\mathcal{M}_{n,1}(\mathbb{R})$. Par conséquent,

Le noyau et l'image de Σ_Y sont supplémentaires orthogonaux dans $\mathcal{M}_{n,1}(\mathbb{R})$.

26 Soit $j \in \llbracket 1; d \rrbracket$. Appliquons le résultat de la question 23 au vecteur $U = V_j$ et à la variable aléatoire $Y - \mathbb{E}(Y)$:

$$\mathbb{V} \left(V_j^\top (Y - \mathbb{E}(Y)) \right) = V_j^\top \Sigma_{Y - \mathbb{E}(Y)} V_j$$

Le vecteur $\mathbb{E}(Y)$ étant constant, le résultat de la question 16 donne alors

$$\Sigma_{Y - \mathbb{E}(Y)} = \Sigma_Y$$

d'où

$$\mathbb{V} \left(V_j^\top (Y - \mathbb{E}(Y)) \right) = V_j^\top \Sigma_Y V_j$$

Enfin, comme $V_j \in \text{Ker } \Sigma_Y$, on a $\Sigma_Y V_j = 0$. Finalement,

$$\boxed{\forall j \in \llbracket 1; d \rrbracket \quad \mathbb{V} \left(V_j^\top (Y - \mathbb{E}(Y)) \right) = 0}$$

27 Soit $j \in \llbracket 1; d \rrbracket$. D'après le résultat de la question 26, la variance de la variable aléatoire $V_j^\top (Y - \mathbb{E}(Y))$ est nulle. Cette variable est à valeurs dans $\mathbb{R}$, donc d'après le cours, cette variable est presque sûrement constante égale à son espérance. Notons $V_j = (v_{j,1} \ \dots \ v_{j,n})^\top$ et calculons l'espérance de la variable $V_j^\top (Y - \mathbb{E}(Y))$, en utilisant la linéarité de l'espérance :

$$\mathbb{E} \left(V_j^\top (Y - \mathbb{E}(Y)) \right) = \mathbb{E} \left(\sum_{i=1}^n v_{j,i} (Y_i - \mathbb{E}(Y_i)) \right) = \sum_{i=1}^n v_{j,i} \mathbb{E}(Y_i - \mathbb{E}(Y_i)) = 0$$

On en déduit que $V_j^\top (Y - \mathbb{E}(Y))$ est presque sûrement nulle, ce qui s'écrit

$$\boxed{\forall j \in \llbracket 1; d \rrbracket \quad \mathbb{P} \left(V_j^\top (Y - \mathbb{E}(Y)) = 0 \right) = 1}$$

28 En utilisant les notations de l'énoncé, le résultat de la question 25 implique que

$$(Y - \mathbb{E}(Y) \in \text{Im } \Sigma_Y) = (Y - \mathbb{E}(Y) \in (\text{Ker } \Sigma_Y)^\perp)$$

La famille $(V_1, \dots, V_d)$ est une base orthonormée de $\text{Ker } \Sigma_Y = \text{Vect}(V_1, \dots, V_d)$. Rappelons qu'un vecteur est orthogonal à $\text{Ker } \Sigma_Y$ si, et seulement si, il est orthogonal à chacun des vecteurs de cette base, ce qui s'écrit

$$\text{Im } \Sigma_Y = \text{Ker } \Sigma_Y^\perp = \bigcap_{j=1}^d V_j^\perp$$

On en déduit que

$$(Y - \mathbb{E}(Y) \in \text{Im } \Sigma_Y) = \bigcap_{j=1}^d (V_j^\top (Y - \mathbb{E}(Y)) = 0)$$

Attention à ne pas confondre les deux notations d'orthogonal $V_j^\perp$ (qui est un sous-espace vectoriel) et de transposée $V_j^\top$ (qui est un vecteur ligne).

D'après le résultat de la question 27, la probabilité de chaque événement de cette intersection est égale à 1. D'après le cours, une intersection d'événements presque sûrs est presque sûre. Par conséquent,

$$\text{Si } r < n, \text{ alors } \mathbb{P}(Y - \mathbb{E}(Y) \in \text{Im } \Sigma_Y) = 1.$$

Finalement, au vu du résultat de la question 24, on a dans tous les cas

$$\boxed{\mathbb{P}(Y - \mathbb{E}(Y) \in \text{Im } \Sigma_Y) = 1}$$

III. MAXIMISATION DE LA VARIANCE

29 La matrice A_2 est diagonale et tous ses coefficients diagonaux (donc ses valeurs propres) sont positifs. En appliquant la question 21,

Il existe un vecteur aléatoire dont A_2 est la matrice de covariance.

30 Considérons une variable aléatoire Y à valeurs dans $\mathcal{M}_{3,1}(\mathbb{R})$ telle que $\Sigma_Y = A_2$. Soit $U \in \mathcal{M}_{3,1}(\mathbb{R})$. D'après le résultat de la question 23,

$$q_Y(U) = \mathbb{V}(U^\top Y) = U^\top \Sigma_Y U = U^\top A_2 U$$

Comme la matrice A_2 est symétrique et que ses valeurs propres 9, 5 et 4 sont positives, le résultat de la question 14 assure alors que

$$\max_{U \in C} U^\top A_2 U = \rho(A_2) = 9$$

Si une variable aléatoire Y vérifie $\Sigma_Y = A_2$, alors $\max_{U \in C} q_Y(U) = \rho(A_2) = 9$.

31 Comme la matrice Σ_Y est symétrique et que ses valeurs propres sont positives d'après les questions 16 et 19, le résultat de la question 23 assure que

$$\forall U \in C \quad q_Y(U) = \mathbb{V}(U^\top Y) = U^\top \Sigma_Y U$$

D'après le résultat de la question 14, la fonction q_Y admet un maximum sur C , égal à $\rho(\Sigma_Y) = \max(\text{sp}(\Sigma_Y))$. Par suite,

La fonction q_Y admet un maximum sur C , égal à $\max(\text{sp}(\Sigma_Y))$.

Notons λ_0 la plus grande valeur propre de Σ_Y et U un vecteur propre associé. En notant $U_0 = U/\|U\|$, on a $U_0 \in C$ et U_0 est également un vecteur propre associé à λ_0 . Le calcul effectué à la question 12 montre alors que

$$q_Y(U_0) = U_0^\top \Sigma_Y U_0 = \lambda_0 = \max(\text{sp}(\Sigma_Y))$$

En prenant $U_0 \in C$ un vecteur propre de Σ_Y associé à sa plus grande valeur propre λ_0 ,

$$\max_{U \in C} \mathbb{V}(U^\top Y) = \mathbb{V}(U_0^\top Y)$$

32 Soit $(i, j) \in \llbracket 1; n \rrbracket^2$ avec $i \neq j$. Par définition,

$$\sigma_{i,j} = \text{cov}(Y_i, Y_j) = \mathbb{E}[(Y_i - \mathbb{E}(Y_i))(Y_j - \mathbb{E}(Y_j))]$$

Or, d'après le cours, quelles que soient les variables aléatoires X et Y possédant une variance, on a

$$\mathbb{E}(XY)^2 \leq \mathbb{E}(X^2)\mathbb{E}(Y^2)$$

Appliquons cette inégalité à $X = Y_i - \mathbb{E}(Y_i)$ et $Y = Y_j - \mathbb{E}(Y_j)$, qui possèdent bien une variance comme Y_i et Y_j . Il vient

$$\mathbb{E}[(Y_i - \mathbb{E}(Y_i))(Y_j - \mathbb{E}(Y_j))]^2 \leq \mathbb{E}(Y_i - \mathbb{E}(Y_i))^2 \mathbb{E}(Y_j - \mathbb{E}(Y_j))^2$$

soit

$$\text{cov}(Y_i, Y_j)^2 \leq \mathbb{V}(Y_i)\mathbb{V}(Y_j)$$

ce qui s'écrit

$$(\sigma^2 \gamma)^2 \leq \sigma^2 \sigma^2$$

d'où $\gamma^2 \leq 1$

puisque $\sigma > 0$. Comme $\gamma > 0$, il vient

$$\boxed{\gamma \leq 1}$$

En outre, en notant I_n la matrice identité de $\mathcal{M}_n(\mathbb{R})$, on a par construction

$$\boxed{\Sigma_Y = \sigma^2 \gamma J + \sigma^2(1 - \gamma)I_n}$$

33 La matrice J est symétrique réelle, donc elle est diagonalisable. Puisque son rang est égal à 1 et que $n \geq 2$, elle admet 0 pour valeur propre et le sous-espace propre associé est de dimension $n - 1$ d'après le théorème du rang. Comme sa trace est égale à n et qu'elle est diagonalisable, la somme de ses valeurs propres est égale à n . Ainsi, la deuxième valeur propre est égale à n et le sous-espace propre associé est de dimension 1. Puisque la somme des colonnes de J vaut $(n \ \dots \ n)^\top$, on en déduit que le vecteur $V = (1 \ \dots \ 1)^\top$ est un vecteur propre associé à la valeur propre n . Finalement,

Les valeurs propres de J sont 0 et n et les sous-espaces propres associés sont respectivement de dimensions $n - 1$ et 1. Le vecteur $V = (1 \ \dots \ 1)^\top$ est un vecteur propre associé à la valeur propre n .

Dans le cas où une matrice J est de rang 1, la base du sous-espace propre associé à la valeur propre non nulle est également une base de $\text{Im } J$, c'est pourquoi l'on peut prendre une colonne de J comme vecteur propre.

34 D'après le résultat de la question 31, le maximum de la variance de Z existe, il est égal à la plus grande valeur propre de Σ_Y et il est atteint pour $U_0 \in C$ vecteur propre associé. Or, d'après le résultat de la question 32,

$$\Sigma_Y = \sigma^2 \gamma J + \sigma^2(1 - \gamma)I_n$$

Comme les vecteurs propres de J sont évidemment des vecteurs propres de I_n , alors les matrices J et I_n admettent une base commune de vecteurs propres. Ce sont des vecteurs propres de toute combinaison linéaire de J et de I_n . Par suite, d'après l'étude réalisée à la question 33, une base de vecteurs propres de $\Sigma_Y = \sigma^2 \gamma J + \sigma^2(1 - \gamma)I_n$ est constituée d'une base de $\text{Ker } J$ (de dimension $n - 1$) et du vecteur $V = (1 \ \dots \ 1)^\top$. Pour déterminer les valeurs propres de Σ_Y ,

- Si $U \in \text{Ker } J$, alors

$$\Sigma_Y U = \sigma^2 \gamma J U + \sigma^2(1 - \gamma)U = \sigma^2(1 - \gamma)U$$

donc $\sigma^2(1 - \gamma)$ est valeur propre de Σ_Y de multiplicité $n - 1$.

- Comme $V = (1 \ \dots \ 1)^\top$ est un vecteur propre de J associé à la valeur propre n , alors

$$\Sigma_Y V = \sigma^2 \gamma J V + \sigma^2(1 - \gamma)V = \sigma^2 \gamma n V + \sigma^2(1 - \gamma)V = \sigma^2(1 + (n - 1)\gamma)V$$

donc $\sigma^2(1 + (n - 1)\gamma)$ est valeur propre de Σ_Y . Sa multiplicité vaut 1 puisque l'hypothèse $n \geq 2$ implique que $\sigma^2(1 + (n - 1)\gamma) > \sigma^2(1 - \gamma)$.

On pouvait également déterminer les valeurs et vecteurs propres de Σ_Y de la façon suivante. Soit λ une valeur propre de Σ_Y et $U \in \mathbb{C}$ un vecteur propre associé. Alors, d'une part, $\Sigma_Y U = \lambda U$ et, d'autre part,

$$\Sigma_Y U = \sigma^2 \gamma J U + \sigma^2 (1 - \gamma) U$$

d'où

$$\lambda U = \sigma^2 \gamma J U + \sigma^2 (1 - \gamma) U$$

puis

$$J U = \frac{\lambda - \sigma^2 (1 - \gamma)}{\sigma^2 \gamma} U \quad \text{car } \sigma > 0 \text{ et } \gamma > 0$$

D'après l'étude réalisée à la question 33, il vient

$$\frac{\lambda - \sigma^2 (1 - \gamma)}{\sigma^2 \gamma} \in \text{sp}(J) = \{0, n\}$$

$$\text{Or, } \frac{\lambda - \sigma^2 (1 - \gamma)}{\sigma^2 \gamma} = 0 \iff \lambda = \sigma^2 (1 - \gamma)$$

$$\text{et } \frac{\lambda - \sigma^2 (1 - \gamma)}{\sigma^2 \gamma} = n \iff \lambda = \sigma^2 + (n - 1) \sigma^2 \gamma$$

$$\text{donc } \lambda \in \{\sigma^2 (1 - \gamma), \sigma^2 + (n - 1) \sigma^2 \gamma\}$$

Réciproquement, soit U un vecteur non nul de $\text{Ker}(J)$ (il existe car la dimension de cet espace vaut $n - 1 \geq 1$ d'après la question 33). Alors

$$\Sigma_Y U = \sigma^2 \gamma J U + \sigma^2 (1 - \gamma) I_n U = \sigma^2 (1 - \gamma) U$$

donc $\sigma^2 (1 - \gamma)$ est une valeur propre de Σ_Y . De plus, en notant V le vecteur déterminé à la question 33, c'est un vecteur propre (donc non nul) pour J associé à la valeur propre n , et

$$\Sigma_Y V = \sigma^2 \gamma J V + \sigma^2 (1 - \gamma) I_n V = \sigma^2 \gamma n V + \sigma^2 (1 - \gamma) V = \sigma^2 (1 + (n - 1) \gamma) V$$

donc $\sigma^2 (1 + (n - 1) \gamma)$ est également une valeur propre de Σ_Y . Par suite,

$$\text{sp}(\Sigma_Y) = \{\sigma^2 (1 - \gamma), \sigma^2 (1 + (n - 1) \gamma)\}$$

Pour appliquer la question 31, notons que la plus grande valeur propre de Σ_Y est égale à $\sigma^2 (1 + (n - 1) \gamma)$ et qu'un vecteur propre associé est le vecteur $V = (1 \ \dots \ 1)^\top$ de norme $\|V\| = \sqrt{n}$. Par conséquent,

La variance de $Z = U_0^\top Y$ est maximale pour $U_0 = \frac{1}{\sqrt{n}} (1 \ \dots \ 1)^\top$.

35 D'après le résultat de la question 31 et le calcul effectué à la question 34, on a, pour la valeur de U_0 déterminée à la question 34,

$$\mathbb{V}(Z) = \rho(\Sigma_Y) = \sigma^2 + (n - 1) \sigma^2 \gamma$$

En outre,
$$\mathbb{V}_T(Y) = \sum_{i=1}^n \mathbb{V}(Y_i) = \text{Tr } \Sigma_Y = n \sigma^2$$

donc

$$\frac{\mathbb{V}(Z)}{\mathbb{V}_T(Y)} = \frac{\sigma^2 + (n - 1) \sigma^2 \gamma}{n \sigma^2} = \frac{1 + (n - 1) \gamma}{n}$$

Comme la variance peut être interprétée comme la quantité d'information contenue dans une variable aléatoire, ce quotient représente la proportion d'information conservée en ne considérant que la première composante principale U_0 .

36 Remarquons tout d'abord que $U_0^\perp$ est un sous-espace vectoriel de dimension non nulle de l'espace de dimension finie $\mathcal{M}_{n,1}(\mathbb{R})$, il est donc fermé. Comme C est également fermé, que $C' = C \cap U_0^\perp$ et que l'intersection de deux fermés est un fermé, on en déduit que l'espace C' est une partie fermée de $\mathcal{M}_{n,1}(\mathbb{R})$.

En outre, la partie $C' \subset C$ est également bornée dans $\mathcal{M}_{n,1}(\mathbb{R})$ puisque C l'est. Puisque l'orthogonal de U_0 est de dimension supérieure ou égale à 1, il contient donc des vecteurs unitaires, ce qui implique que C' est non vide. Comme l'application q_Y est continue, elle est donc bornée sur C' et elle atteint ses bornes sur C' .

L'application q_Y admet un maximum sur C' .

37 Comme la matrice Σ_Y est symétrique, la question 1 assure qu'elle est orthodagonalisable. Notons $(V_1, \dots, V_n)$ une base orthonormée de vecteurs propres de Σ_Y , associés aux valeurs propres $(\lambda_1, \dots, \lambda_n)$. D'après le résultat de la question 31, on peut prendre $V_1 = U_0$, qui est un vecteur propre associé à la plus grande valeur propre de Σ_Y , ici λ_1 .

Soit $U \in C'$. Comme $V_1 = U_0$ et que $C' = C \cap U_0^\perp$, on a $C' \subset \text{Vect}(V_2, \dots, V_n)$, donc on peut écrire

$$U = \sum_{i=2}^n u_i V_i$$

avec $u_i \in \mathbb{R}$ pour tout $i \in \llbracket 2; n \rrbracket$. Ainsi,

$$q_Y(U) = \mathbb{V}(U^\top Y) = U^\top \Sigma_Y U \quad (\text{question 23})$$

$$\begin{aligned} &= \left(\sum_{i=2}^n u_i V_i \right)^\top \Sigma_Y \left(\sum_{j=2}^n u_j V_j \right) \\ &= \sum_{i=2}^n u_i V_i^\top \sum_{j=2}^n u_j (\Sigma_Y V_j) \\ &= \sum_{i=2}^n u_i V_i^\top \sum_{j=2}^n u_j (\lambda_j V_j) \\ &= \sum_{i=2}^n \sum_{j=2}^n \lambda_j u_i u_j (V_i^\top V_j) \end{aligned}$$

d'où

$$q_Y(U) = \sum_{i=2}^n \lambda_i u_i^2$$

puisque la base $(V_1, \dots, V_n)$ est orthogonale. Il vient

$$q_Y(U) \leq \lambda_2 \underbrace{\sum_{i=2}^n u_i^2}_{\|U\|^2=1} = \lambda_2$$

puisque le vecteur U est unitaire et que $\text{Vect}(V_2, \dots, V_n)$ est une famille orthonormée. En outre, la valeur λ_2 est atteinte par $q_Y(U)$ pour $U = V_2$. Par suite,

Le maximum de q_Y sur C' est égal à la deuxième plus grande valeur propre λ_2 de Σ_Y et il est atteint en $U_1 \in C'$ vecteur propre unitaire de Σ_Y associé à λ_2 .

Remarquons que l'on applique simplement dans cette question les résultats des parties précédentes à l'endomorphisme induit par Σ_Y sur $U_0^\perp$.

38 D'après le cours, comme $U_0^\top Y$ et $U_1^\top Y$ sont des variables aléatoires réelles,

$$\text{cov}(U_0^\top Y, U_1^\top Y) = \frac{1}{4} \left(\mathbb{V}((U_0 + U_1)^\top Y) - \mathbb{V}((U_0 - U_1)^\top Y) \right)$$

On peut le redémontrer facilement par le calcul suivant :

$$\begin{aligned} & \mathbb{V}(U_0^\top Y + U_1^\top Y) = \mathbb{V}(U_0^\top Y) + \mathbb{V}(U_1^\top Y) + 2 \text{cov}(U_0^\top Y, U_1^\top Y) \\ \text{et } & \mathbb{V}(U_0^\top Y - U_1^\top Y) = \mathbb{V}(U_0^\top Y) + \mathbb{V}(U_1^\top Y) - 2 \text{cov}(U_0^\top Y, U_1^\top Y) \\ \text{donc } & \text{cov}(U_0^\top Y, U_1^\top Y) = \frac{1}{4} \left(\mathbb{V}(U_0^\top Y + U_1^\top Y) - \mathbb{V}(U_0^\top Y - U_1^\top Y) \right) \\ & = \frac{1}{4} \left(\mathbb{V}((U_0 + U_1)^\top Y) - \mathbb{V}((U_0 - U_1)^\top Y) \right) \end{aligned}$$

Le résultat de la question 23 assure que

$$\begin{cases} \mathbb{V}((U_0 + U_1)^\top Y) = (U_0 + U_1)^\top \Sigma_Y (U_0 + U_1) \\ \mathbb{V}((U_0 - U_1)^\top Y) = (U_0 - U_1)^\top \Sigma_Y (U_0 - U_1) \end{cases}$$

En outre, d'après les questions 31 et 37,

$$\begin{cases} \Sigma_Y U_0 = \lambda_1 U_0 \\ \Sigma_Y U_1 = \lambda_2 U_1 \\ U_0^\top U_0 = U_1^\top U_1 = 1 \\ U_0^\top U_1 = U_1^\top U_0 = 0 \end{cases}$$

$$\begin{aligned} \text{Il vient } & \mathbb{V}((U_0 + U_1)^\top Y) = (U_0 + U_1)^\top \Sigma_Y (U_0 + U_1) \\ & = (U_0 + U_1)^\top (\lambda_1 U_0 + \lambda_2 U_1) \\ & = \lambda_1 U_0^\top U_0 + \lambda_1 U_1^\top U_0 + \lambda_2 U_0^\top U_1 + \lambda_2 U_1^\top U_1 \end{aligned}$$

$$\mathbb{V}((U_0 + U_1)^\top Y) = \lambda_1 + \lambda_2$$

$$\begin{aligned} \text{et } & \mathbb{V}((U_0 - U_1)^\top Y) = (U_0 - U_1)^\top \Sigma_Y (U_0 - U_1) \\ & = (U_0 - U_1)^\top (\lambda_1 U_0 - \lambda_2 U_1) \\ & = \lambda_1 U_0^\top U_0 - \lambda_1 U_1^\top U_0 - \lambda_2 U_0^\top U_1 + \lambda_2 U_1^\top U_1 \end{aligned}$$

$$\mathbb{V}((U_0 - U_1)^\top Y) = \lambda_1 + \lambda_2$$

Finalement,

$$\begin{aligned} \text{cov}(U_0^\top Y, U_1^\top Y) & = \frac{1}{4} \left(\mathbb{V}((U_0 + U_1)^\top Y) - \mathbb{V}((U_0 - U_1)^\top Y) \right) \\ & = \frac{1}{4} (\lambda_1 + \lambda_2 - (\lambda_1 + \lambda_2)) \end{aligned}$$

d'où

$$\boxed{\text{cov}(U_0^\top Y, U_1^\top Y) = 0}$$

Ceci signifie que les deux premières composantes principales sont décorréliées entre elles.